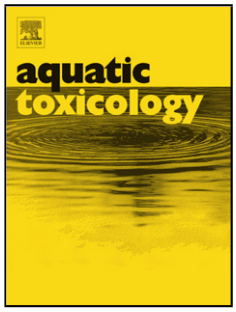
Key metabolic pathways involved in xenobiotic biotransformation and stress responses revealed by transcriptomics of the mangrove oyster Crassostrea brasiliana

Lüchmann, Karim H.; Clark, Melody S.; Bainy, Afonso C.d.; Gilbert, Jack A.; Craft, John A.; Chipman, J.Kevin; Thorne, Michael A.s.; Mattos, Jacó J.; Siebert, Marília N.; Schroeder, Declan C.

DOI:
10.1016/j.aquatox.2015.06.012



Title: Key metabolic pathways involved in xenobiotic biotransformation and stress responses revealed by transcriptomics of the mangrove oyster *Crassostrea brasiliana*

Author: Karim H. Lüchmann Melody S. Clark Afonso C.D. Bainy Jack A. Gilbert John A. Craft J.Kevin Chipman Michael A.S. Thorne Jacó J. Mattos Marília N. Siebert Declan C. Schroeder

PII: S0166-445X(15)30003-5
DOI: http://dx.doi.org/doi:10.1016/j.aquatox.2015.06.012
Reference: AQTOX 4148

To appear in: *Aquatic Toxicology*

Received date: 31-3-2015
Revised date: 25-6-2015
Accepted date: 26-6-2015

Please cite this article as: Lüchmann, Karim H., Clark, Melody S., Bainy, Afonso C.D., Gilbert, Jack A., Craft, John A., Chipman, J.Kevin, Thorne, Michael A.S., Mattos, Jacó J., Siebert, Marília N., Schroeder, Declan C., Key metabolic pathways involved in xenobiotic biotransformation and stress responses revealed by transcriptomics of the mangrove oyster Crassostrea brasiliana.Aquatic Toxicology http://dx.doi.org/10.1016/j.aquatox.2015.06.012

Key metabolic pathways involved in xenobiotic biotransformation and stress responses revealed by transcriptomics of the mangrove oyster Crassostrea brasiliana

Karim H. Lüchmann[a]*, Melody S. Clark[b], Afonso C.D. Bainy[c], Jack A. Gilbert[d,e,f,g], John A. Craft[h], J. Kevin Chipman[i], Michael A.S. Thorne[b], Jacó J. Mattos[c], Marília N. Siebert[c], Declan C. Schroeder[j]*

[a] Fishery Engineering Department, Santa Catarina State University, Laguna, Brazil

[b] British Antarctic Survey, Natural Environment Research Council, Cambridge, UK

[c] Biochemistry Department, Federal University of Santa Catarina, Florianópolis, Brazil

[d] Biosciences Division (BIO), Argonne National Laboratory, Argonne, USA

[e] Department of Ecology and Evolution, University of Chicago, Chicago, USA

[f] Marine Biological Laboratory, Woods Hole, USA

[g] College of Environmental and Resource Sciences, Zhejiang University, Hangzhou, China

[h] Biological and Biomedical Sciences, Glasgow Caledonian University, Glasgow, UK

[i] School of Biological Sciences, The University of Birmingham, Birmingham, UK

[j] Marine Biological Association of the United Kingdom (MBA), Plymouth, UK

***** Corresponding authors.

E-mail address of each author: khluchmann@gmail.com, karim.luchmann@udesc.br (KHL); mscl@bas.ac.uk (MSC); afonso.bainy@ufsc.br (ACDB); gilbertjack@anl.gov (JAG); J.A.Craft@gcu.ac.uk (JAC); j.k.chipman@bham.ac.uk (JKC); mior@bas.ac.uk (MAST); jaco.mattos@ufsc.br (JJM); marilia.siebert@ifsc.edu.br (MNS); dsch@mba.ac.uk (DCS)

Highlights

- Thetranscriptome of the mangrove oyster *Crassostrea brasiliana*was sequenced.
- The sequencing effort dramatically expanded the *existing cDNA sequences available for the species.*
- Global analysis for transcription in the oyster treated with phenanthrene, diesel and domestic sewage was performed.
- The pollutants altered mRNAs for genes in biotransformation, antioxidant and stress response pathways.

Abstract

The Brazilian oyster *Crassostrea brasiliana* was challenged to three common environmental contaminants: phenanthrene, diesel fuel water-accommodated fraction (WAF) and domestic sewage. Total RNA was extracted from the gill and digestive gland, and cDNA libraries were sequenced using the 454 FLX platform. The assembled transcriptome resulted in ~20,000 contigs, which were annotated to produce the first *de novo* transcriptome for *C. brasiliana*. Sequences were screened to identify genes potentially involved in the biotransformation of xenobiotics and associated antioxidant defence mechanisms. These gene families included those of the cytochrome P450 (CYP450), 70kDa heat shock, antioxidants, such as glutathione *S*-transferase, superoxide dismutase, catalase and also multi-drug resistance proteins. Analysis showed that the massive expansion of the CYP450 and HSP70 family due to gene duplication identified in the *Crassostrea gigas* genome also occurred in *C. brasiliana*, suggesting these processes form the base of the *Crassostrea* lineage. Preliminary expression analyses revealed several candidates biomarker genes that were up-regulated during each of the three treatments, suggesting the potential for environmental monitoring.

Keywords: xenobiotic metabolism; antioxidant parameters; pollutants; bioaccumulation; bivalve; polycyclic aromatic hydrocarbon.

1. Introduction

Mangrove oysters, *Crassostrea brasiliana* (sin. *Crassostrea gasar*, Lazoski et al., 2011), are common species along the Brazilian coast, where they are both economically and ecologically important. These sessile, filter feeders are known to accumulate water contaminants in their tissues and are therefore ideal bioindicator species for pollution monitoring in coastal waters (Lüchmann et al., 2011; 2014). However, along with other Ostreidae species, little is known regarding the specific genomic and transcriptomic adaptations to these contaminants. The monitoring endpoints are based on a small number of biomarkers with origins in human toxicology, and are therefore not bivalve-specific (Forbes et al., 2006). In excess of the ecological and economical importance of oysters to the coastal areas and the aquaculture industry, understanding their biology, susceptibility to pollutants and differential stress resistance has become an important issue for modern ecotoxicology. In particular, genomic resources such as genome or transcriptome sequences would greatly facilitate studies into the cellular

mechanisms under-pinning biological responses in this species and enable the development of molecular markers for bioaccumulative pollution monitoring.

Due to their economic importance, oysters have been the subject of several large-scale expressed sequence tags (EST) projects (Fleury et al., 2009; Joubert et al., 2010; Tanguy et al., 2008; Wang and Guo, 2007). Indeed whole genome or transcriptome sequencing has proved a very efficient and cost effective method for expanding the sequence database for bivalves and other non-model species (*i.e.* Clark et al., 2010; Craft et al., 2010; Hou et al., 2011; Joubert et al., 2010; Meyer et al., 2009), and so was applied here to the transcriptome of the mangrove oyster *C. brasiliana.* In the last few years, several oyster genomes have become available, including the Pacific oyster *Crassostrea gigas* (Zhang et al., 2012a) and Pearl oyster *Pinctada fucata* (Takeuchi et al., 2012), significantly enriching the genomic resources for this animal model.

Here we aimed to sequence the transcriptome of the mangrove oyster *C. brasiliana* to both improve the genomic resources for this species, and to explore gene transcription for biotransformation of xenobiotics, antioxidant and stress response during exposure to three different environmental contaminants: phenanthrene (PHE), diesel fuel water-accommodated fraction (diesel WAF) and domestic sewage. All three contaminants are key chemical models for ecotoxicological studies. Phenanthrene, a 3-ring compound included in the US-EPA priority pollutant list, is one of the most abundant aquatic PAH (polycyclic aromatic hydrocarbon), as a result of human activities (US EPA, 2009). It is lipophilic and has a low molecular weight, making it easily taken-up by aquatic organisms (Oliveira et al., 2007), with a greater bioaccumulation rate in bivalve molluscs (Hannam et al., 2010; Lüchmann et al., 2014). In contrast, diesel WAF comprises a model for complex mixtures derived from petroleum industry activities. Diesel fuel is one of the most common aquatic contaminants, and has recently been shown to exert biochemical effects and bioaccumulation trends in *C. brasiliana* (Lüchmann et al., 2011). Domestic sewage was chosen based on the high inputs of untreated sewage discharges in coastal ecosystems around the world and its potential effects on transcriptional levels of oysters (Medeiros et al., 2008).

C. brasiliana was challenged to each contaminant separately, total RNA was extracted from the gill and digestive gland with the resulting cDNA libraries sequenced using the 454 FLX platform. The sequence data was assembled into a reference transcriptome, which was then screened to identify genes potentially involved in the biotransformation of xenobiotics and associated antioxidant defence and stress mechanisms. The results demonstrated differences between the

responses to the different toxicants, with promising relevance for ecotoxicology studies and aquatic monitoring programs.

2. Material and methods

2.1. Oyster collection and chemical exposures

Mangrove oysters (*Crassostrea brasiliana*) of similar shell length (5-8 cm) were collected from an oyster farm at Sambaqui beach (Marine Mollusks Laboratory, UFSC) in Florianópolis, southern Brazil. This criterion was strictly adhered to and therefore limited availability of oysters throughout the course of the study meant that certain experiments were unfortunately run without replication. After collection, the animals were covered with wet towels and immediately transported in coolers by road approximately 20 km to the laboratory. In the lab, the oysters were transferred into 50 litre aquaria containing 0.45 μm-filtered, aerated seawater, at 21 °C, and salinity 25. Oysters were fed twice a day on microalgae (*Chaetoceros muelleri* and *Isochrysis* sp.) at a density of 3.3 x 10^6 cels mL^{-1} and 2.2 x 10^6 cels mL^{-1}, respectively, and water was changed daily for one week prior to experiments. Oyster were then randomly divided into the glass exposure tanks (1 animal per 1 L of seawater) and held (without feeding) for 24 h prior to the exposures. During the exposure periods, control and exposed organisms were not fed to prevent potential bioaccumulation of chemicals by food.

There were 4 exposure experiments: diesel WAF for 24 h, diesel WAF for 72 h, PHE for 24 h and sewage for 24 h, which were carried out in different occasions but the oysters were supplied from the same brood stock of the mollusc farm, and were submitted to the same acclimatization process as described above. For each set of experiment, there was a control group where a separate set of oysters was kept under control conditions in normal seawater, with the exception of the PHE exposure control group, where the seawater also included 0.01% DMSO, as this was the solvent used to dissolve the PHE (please see details of the exposure condition below). The diesel WAF exposure was carried out in duplicate, and PHE and sewage exposures were performed without replication.

Diesel fuel was purchased at a PETROBRAS petrol station and WAF was obtained according to Singer et al. (2000) with minor modifications. Briefly, one part (1 L) of fresh diesel fuel was diluted with nine parts (9 L) of 0.45 μm-filtered seawater (salinity 25) in a sealed 14 L glass flask which was protected from light, in order to minimize evaporation and degradation of the fuel components. The diesel-water mixture was stirred for 23 h with the homogenizer Glas-Col (LLC)

using a steel modified pestle at 1600 rpm at a constant temperature of 21 °C. The mixture was then allowed to settle for 1 h before the lower layer of water (diesel WAF) was transferred into the glass tanks. The 10% diesel WAF was prepared through dilution of the WAF with the control seawater. The diesel WAF concentration was chosen based on previous results of biochemical biomarkers measured in *C. brasiliana* (Lüchmann et al., 2011). No mortality was observed in the control and treated groups. The levels of individual and total PAHs bioaccumulated after 24 h are summarized in the Supplementary Table S1.

Phenanthrene (PHE) (Sigma-Aldrich, P1, 140-9) was first dissolved in dimethyl sulfoxide (DMSO), and then added to 0.45 μm-filtered seawater (salinity 25) to achieve final nominal PHE concentration of 1000 $\mu g.L^{-1}$ (equivalent to 5.6 μM), and a final maximum DMSO concentration of 0.01% (v/v). The concentration of PHE added to the test media was chosen based on previous reports carried out with bivalves (*i.e.* Hannam et al., 2010; Lüchmann et al., 2011; 2014). Oysters were then randomly divided into the glass exposure tanks, which were individually aerated and covered with glass and sealed to avoid evaporation of PHE, and held (without feeding) for 24 h prior to the exposure. The control oysters were subjected to the same conditions as the exposed groups, except for the addition of 0.01% (v/v) DMSO only without PHE. No mortality was observed in the control and treated groups.

Sewage exposure was performed according to Medeiros et al. (2008) with minor modifications. Briefly, domestic sewage was collected at the influent duct of the downtown wastewater treatment plant (Florianópolis, southern Brazil) after solid material grid removal, and diluted to 33% (v/v) using 0.45 μm-filtered seawater (salinity 25). Oysters were placed in the exposure glass tanks which were individually aerated using glass Pasteur pipettes and were covered with glass. No mortality was observed in the control and treated groups.

After chemical exposures, twelve oysters from each of the diesel fuel WAF and PHE experiments, and seven from the domestic sewage experiment were sacrificed and the gill and digestive gland were immediately excised, flash frozen in liquid nitrogen and individually stored at -80 °C until further analysis. Three oysters from the control groups of each treatment were pooled, totalizing 12 animals for the control. The control group included oysters from both seawater control and DMSO-control.

2.2. Total RNA isolation and preparation of cDNA libraries

Total RNA from the gill and digestive gland of each oyster was individually isolated using TRIzol reagent (Invitrogen, UK) and purified with the Nucleospin RNA II Total RNA Isolation Kit (AbGene, UK) following the supplier's protocol with minor modifications. Briefly, 50 – 100mg of each tissue was mechanically disrupted in the presence of 1mL TRIzol using a homogenizer (Tissue-Tearor, BioSpec Products). The TRIzol protocol was strictly followed until separation of the phases, with 200µL of the upper aqueous phase transferred to a new tube for on-column purification using the Nucleospin RNA II Total RNA Isolation Kit (AbGene, UK). RNA was finally eluted in 60µL of RNase-free water. Residual genomic DNA contamination was removed during the RNA Nucleospin cleanup using the DNase I digestion following manufacturer's instructions (AbGene, UK). The integrity of the purified total RNA was assessed using formaldehyde agarose gel electrophoresis, and RNA quantity was determined using a NanoDrop ND-1000 spectrophotometer (ThermoScientific, UK).

Equal quantities of purified total RNA (3µg) were pooled into 10 samples: gill and digestive gland for each of diesel WAF 24 h, diesel WAF 72 h, PHE, sewage and control groups. The pools were used for synthesis of non-normalized full-length double-stranded cDNA. cDNA libraries were constructed for each sample using the SMARTer PCR cDNA Synthesis Kit (Clontech, Paris) according to the manufacturer's instructions. Full-length cDNA templates were then amplified by long-distance PCR using the Advantage 2 PCR Kit (Clontech, Paris). To ensure that the PCR products were not over amplified, the optimal number of PCR cycles was determined according to the manufacturer's guidelines, which was verified by agarose gel electrophoresis. The products were purified with the DNA Clean & Concentrator™-5 Kit (Zymo Research, USA). The amplified cDNA libraries were verified for quality by microcapillary electrophoresis (Agilent Bioanalyzer 2100, Agilent Technologies) and quantified using a NanoDrop ND-1000 spectrophotometer (ThermoScientific, UK).

2.3. Transcriptome pyrosequencing

The 10 cDNA libraries were then combined into 4 samples. Both gill and digestive gland tissue-specific libraries were combined for each of the controls, sewage, PHE and diesel WAF (with both 24 and 72 h samplings). The libraries were then submitted to size-selection in a gel with two libraries prepared: one selected for "larger" fragments (cut from just below the 4kb and sheared by nebulisation down to 500 – 700 bp) and one for "smaller" fragments (cDNA fragments of approximately 700 bp). Molecular Identifier (MID) tags were used to enable

subsequent identification of size-selected portions from each library dataset. The 8 tagged libraries (two for each cDNA library) were then combined and pyrosequenced on the half of a PicoTiter Plate using the 454 Genome Sequencer FLX System according to the manufacturers' instructions (454/Roche, http://www.454.com). Pyrosequencing was performed by the Centre for Genomic Research at the University of Liverpool.

2.4. Sequence data analysis and assembly

The raw sequence data obtained from the cDNA libraries were pooled and subject to filtering and trimming of SMARTer™ adaptors for cDNA synthesis, primers, poly (A/T) tails and potential contaminating vector sequences. Following the sequence trimming and size selection (>150 bp), the reads were assembled using SeqMan NGen v 3.0.4 (DNASTAR). Default parameters for *de novo* assembly of '454' reads were used, except for two settings: the 'Min Match Percentage' was set to 80 (instead of 85) and minimum number of reads to form a consensus sequences, which was set to a minimum of 2 (instead of 10). The first parameter represents the minimum percentage of identity required to join two sequences in the same contig and the second was chosen in order to assemble low frequency reads, which potentially represented weakly transcribed genes.

2.5. Functional annotation and mapping

The dataset was annotated by first searching the NCBI non-redundant (nr) nucleotide databases using blastn (*E* value threshold of 10^{-6}) to identify rRNA genes and mitochondrial (mtDNA) sequences. The rational for selecting the blastn algorithm for the initial search lies in the absence of blastx matches of most contigs with matches to rRNA genes and mtDNA (data not shown), which could have lead to bias in the types of genes during the annotation of the *C. brasiliana* transcriptome. Sequences with matches to rRNA genes, mtDNA sequences, Prokaryotic and virus proteins, were then excluded from the *C. brasiliana* transcriptome dataset. This final set of contigs provided the reference transcriptome, against which the reads were mapped. The mapping was carried out using the mapping facility of Newbler with default mapping parameters. Both the full and partial mappings to each contig were extracted from all the respective mappings, and the values normalised by total library count. A ratio test was carried out in R, adjusting the p-values for multiple testing (Benjamini and Hochberg, 1995). The ratios from the different treatments relative to the control, or fold changes, are reported in the results.

Both blastn and blastx similarity searches were performed using Blast2GO PRO v.2.5.0 software (Conesa et al., 2005) using a cut-off of 10^{-6}. The resulting top 10 blastx matches were fed into Blast2GO in order to retrieve associated Gene Ontology (GO) terms describing biological processes, molecular functions, and cellular components and extract the corresponding enzyme commission numbers (EC). InterPro terms were also obtained from InterProScan at EBI, converted and merged with GOs using Blast2GO software. Finally, the KEGG (Kyoto Encyclopedia of Genes and Genome) orthology (KO) identifiers, or the K numbers, were generated using the web-based server KAAS (KEGG Automatic Annotation Server) (Moriya et al., 2007), resulting in the mapping of putative KEGG metabolic pathways of *C. brasiliana*. The transcripts assigned to each gene family were first identified based on the closest blastx matches in the NCBI nr database, followed by the identification of at least one of the descriptors provided by the relevant InterPro Scan and Pfam domain (Supplementary Table S2).

3. Results and discussion

3.1. 454 sequencing and *de novo* assembly of the oyster transcriptome

Pooled non-normalized cDNA libraries produced from gill and digestive gland RNA of *C. brasiliana* exposed to phenanthrene (PHE), diesel fuel water-accommodated fraction (WAF) and domestic sewage, plus a control group, were sequenced yielding 399,291 raw reads with an average length of 260 bp (GS FLX 2009 chemistry). In total, 246,514 reads passed quality filtering, while 152,777 reads were identified as singletons and were excluded from further analysis. Assembly of the 246,514 reads produced 20,938 contigs (150 to 4,662 bp), with an average length of 575 bp (Table 1). Annotation identified 282 contigs that matched rRNA, mtDNA, Prokaryotic and virus genes, which were removed from future analysis (Table 1). This final dataset was then subjected to putative functional annotation using a more rigorous blastx search (*E* value < 10^{-10}) against the NCBI non-redundant (nr) protein database. The length of contigs was significantly correlated with the number of sequences assembled into them (Pearson $\rho = 0.53$; $n = 20{,}938$; $p < 0.0001$), as expected for 454 reads (Parchman et al., 2010) and the size distribution of contigs resembled that of previous 454 studies (Bettencourt et al., 2010; Fraser et al., 2011; Hou et al., 2011) (Supplementary Figures S1A and S1B). Files containing the reads have been submitted to the National Center for Biotechnology Information Short Read Archive [GenBank: SRX790900].

Table 1 should be placed here.

3.2. Functional annotation

A total of 47.5% of the contigs were putatively annotated, which was significantly greater than other non-model marine invertebrates in which EST collections have been subjected to pyrosequencing (c.f. 9.9% in the razor clam *Sinonovacula constricta* (Niu et al., 2013), 12% in the blue mussel *Mytilus galloprovincialis* (Craft et al., 2010), 17% in the Antarctic bivalve *Laternula elliptica* (Clark et al., 2010; Meyer et al., 2009), 24% in the Manila clam *Ruditapes philippinarum* (Milan et al., 2011), 25% in the krill *Euphausia superba* (Clark et al., 2011), 28% in the Yesso scallop *Pactinopecten yessoensis* (Hou et al., 2011) and 39% in the sea cucumber *Apostichopus japonicus* (Du et al., 2012). This is almost certainly due to the genome sequencing and annotation of a close relative; the Pacific oyster, *Crassostrea gigas* (Zhang et al., 2012a).

Oyster sequences that had matches in NCBI nr protein database were subject to Gene Ontology (GO) analysis to determine the corresponding enzyme commission (EC) number. EC numbers were assigned to 1,341 sequences, of which, the most important for this study were the oxidoreductases (16.7%) (Figure 1). In addition, 2,673 transcripts were assigned KEGG orthology (KO) identifiers (K numbers), which were then integrated into the KEGG resource by the KAAS service, resulting in the assignment to 276 different KEGG pathways. The number of pathways identified in this study is superior to other previous studies (*i.e.* Hao et al., 2011; Hou et al., 2011) and might suggest that the *C. brasiliana* sequence data contain a large diversity of genes involved in a variety of biological processes, and do not contain notable biases towards particular categories of genes. Of these 276 KEGG pathways, those commonly related to xenobiotic biotransformation and response to stress in aquatic organisms were represented by sequences classified into “metabolism of xenobiotics by cytochrome P450”, “drug metabolism - cytochrome P450”, “drug metabolism - other enzymes”, “glutathione metabolism”, and “pentose phosphate pathway”. These annotations provided a valuable resource for the identification of novel genes involved in the pathways of xenobiotic biotransformation and stress responses.

Figure 1 should be placed here.

3.3. Detection of transcripts encoding genes involved in xenobiotic biotransformation

In Eukaryotes, the biotransformation of most lipophilic xenobiotics can be divided in four phases. In phase 0 there is an initial uptake of the xenobiotic. During phase I, polar groups are introduced into the xenobiotic molecule, which is usually catalysed by the cytochrome P450 enzymes (CYP450), making it a suitable substrate for phase II reactions. Thus, the metabolites formed by phase I reactions may undergo further metabolism by conjugation to polar endogenous substrates which is mainly catalysed by enzymes like glutathione *S*-transferases (GSTs), sulfotransferases (SULT) and glucoronosyltransferases whereby lipophilic compounds are transformed into hydrophilic conjugates (Walker et al., 1996). In phase III specialized transporters recognize the conjugates, and expel them from the cell (Homolya et al., 2003). Therefore, the response of enzymes belonging to these four phases represents a defence mechanism developed by an organism exposed to contaminants and have been used as biomarkers to evaluate both exposure to, and effects of, environmental pollutants. Given this and the importance of oysters for aquatic biomonitoring programs, our aim was to identify transcripts involved in the metabolism of xenobiotic biotransformation and associated antioxidant defence, providing a baseline for future studies on the environmental stress response of *C. brasiliana*. Then, the selection of candidate genes involved in phases I, II and III was based on the database analyses above, using the queries matching xenobiotics detoxification-related criteria to identify relevant sequences from the oyster transcriptome.

3.4. Transcripts encoding putative CYP450s

A total of 43 CYP450-related contigs were identified in the *C. brasiliana* transcriptome. These were assigned to nine CYP families and included representatives of all the major CYP clans in protostomes (2-4, 7, and mitochondrial; Karatolos et al., 2011; Nelson et al., 2013) (Table 2). The majority of transcripts were assigned to the CYP2 family, which is in accordance with data from other marine invertebrates, where this family predominates (Guo et al., 2013; Putnam et al., 2007; Sodergren et al., 2006; Zhang et al., 2012a). Indeed in most species it has been identified as the most diverse CYP family, which has evolved complex roles in physiology and toxicology, and diverse regulatory mechanisms (Kubota et al., 2011). Analysis of the *C. gigas* genome data showed a considerable expansion of genes involved in defence pathways, often called “stress genes”. This included the whole CYP450 gene family with the identification of 136 different genes with over half assigned to the CYP2 family. This represented a considerable increase in the number of CYP genes compared to other species: 57 in human, 78 in Daphnia, 82 in the sea

anemone, 96 in *Capitella teleta* and 120 in the sea urchin and correlation of the transcripts with genome data indicated that this is, in many cases, due to tandem duplications (c.f. Figure S20 in Zhang et al., 2012a). Although such an extensive repertoire of CYP450 genes was not identified in our smaller dataset, this preliminary scan indicates that there may be a similar expansion in *C. brasiliana* highlighting that the origin of such duplications may be at the base of the *Crassostrea* lineage.

Table 2 should be placed here

Among the putative CYP450 families with ecotoxicology interest identified in this study, CYP1 and CYP2 are known for the oxidative transformation of environmental contaminants in deuterostomes (Kubota et al., 2011; Nebert and Russel, 2002). The involvement of CYP families in protostomes is not so well documented, but in insects, the CYP3 and CYP4 families have been implicated in the metabolism of synthetic insecticides (Feyereisen, 2006; Karatolos et al., 2011), and recently the CYP2 and CYP4 families were suggested as being involved in defence against toxic chemicals of bivalve molluscs (Miao et al., 2011; Zanette et al., 2010). Our expression data show that the CYP2 family members (Figure 2; Table 3) are up-regulated in response to both diesel WAF and sewage, with CYP2D expression more specific to diesel WAF, indicating their roles in the detoxification of organic compounds such as PAHs, which has been previously reported in cormorant and mouse (Kubota et al., 2011; Schober et al., 2010). This is consistent with recently documented report that benzo[*a*]pyrene (BaP) induced CYP414A1, a subfamily closely related to members of the CYP2 family, in the clam *Venerupis* (R.) *philippinarum* (Zhang et al., 2012b). Increased levels of CYP2-like mRNA have also been detected in the gill of *C. brasiliana* following short-term exposure to PHE (Lüchmann et al., 2014). Similarly, the CYP4 gene family encodes a diverse number of enzymes, with functions related to the hydroxylation of fatty acids and eicosanoids in vertebrates (Kikuta et al., 2002), and xenobiotic biotransformation in polychaetes and insects (Rewitz et al., 2006). A wide variety of factors have been shown to regulate CYP4 family members, ranging from hormones in vertebrates (Simpson, 1997) to dissolved oxygen in bivalves (Snyder et al., 2001). There is also recent evidence of the regulation by environmental pollutants on CYP4 related genes in bivalves, in which BaP decreased the transcription levels of CYP4 in scallops (Miao et al., 2011). However, the *C. brasiliana* CYP4-like genes did not show any response to diesel WAF, PHE or sewage.

Figure 2 and Table 3 should be placed here

Putative transcripts encoding CYP17 proteins and a closely related family member CYP356A1 were well represented in *C. brasiliana* transcriptome (Table 2). Although CYP17 is thought to be related to the steroid metabolism, the role of homologous sequences of the xenobiotic detoxification is also suggested (Toledo-Silva et al., 2008) and indeed, our data show CYP17 transcripts are up-regulated in response to all the chemical challenges, with the highest responses for PHE and sewage. A similar response was seen in a *CYP356A1-like* gene which was over-expressed in the gill of *C. brasiliana* exposed to PHE (Lüchmann et al., 2014) and together with CYP2 are therefore candidate biomarkers for PAH biotransformation in bivalve molluscs. Furthermore, four sequences could be aligned to the CYP1 family, which is reported to be the main phase I enzyme involved in PAHs metabolism in vertebrates. To date, the role and even the existence of CYP1 in molluscs are poorly described (Zanette et al., 2013), but this study shows that CYP1 is up-regulated in response to diesel WAF and PHE. Finally, while recent studies have shown differential modulation in CYP3 transcription of bivalves (Cubero-Leon et al., 2012; Zanette et al., 2013), our data indicate that *C. brasiliana* CYP3-like genes are induced in a non-specific manner by the chemicals tested in this study.

3.5. Transcripts encoding putative 70kDa heat shock proteins (HSP70)

Heat shock proteins, particularly those of the HSP70 family have long been associated with the response to environmental stress (Reviewed in Gross, 2004). Most members of this family are present in the normal cell state, as they have a housekeeping function with a critical role in the folding of native polypeptides and their translocation to different cellular compartments (Feder and Hofmann, 1999; Hartl and Hartl-Meyer, 2002). However, under conditions which elicit cellular stress, there is an increase in the levels of protein mis-folding. In these cases, heat shock proteins are frequently up-regulated to assist in either protein refolding or targeting the denatured proteins for removal from the cell. The latter is important because if the denatured proteins accumulate in the cell, they are cytotoxic (Fink 1999; Hartl, 1996; Parsell and Lindquist, 1993). 53 contigs were identified with matches to the HSP70 family, which could be assigned to 8 different members of the HSP70 family (Table 4). In the majority of cases, it was possible to assign multiple sequence fragments to a single gene, with 36 fragments mapping to the HSP70-

12 family to closely related members; HSP70-12A and HSP70-12B. Whilst it is always difficult to determine actual gene numbers when dealing with gene fragments, manual searches and alignments indicated that there are at least 13 members of the HSP70-12A family and 8 members of the HSP70-12B family, representing a massive expansion of the HSP70-12A and HSP70-12B families in *C. brasiliana.* These data concur with the analysis of the *C. gigas* genome, where 88 HSP70 genes were identified, with phylogenetic analysis showing that the majority of these are expansions specific to *C. gigas* (Zhang et al., 2012a). These oyster-specific duplication events, both in *C. brasiliana* and *C. gigas* appear to be restricted to the HSP70-12 family (c.f. Figure S19 in Zhang et al., 2012a) and have not affected the other more well known members of the HSP70 family, such as the inducible and constitutive forms (HSP70 and HSC71 respectively). The HSP70-12 genes are atypical HSP70 family members, which are each present in vertebrates in a single copy. They are not classically up-regulated in response to stress, but have been shown to be involved in lipoprotein interactions and act as modulators of inflammation and/or apoptosis in humans (Cui et al. 2010; Han et al., 2003). HSP70-12B has also been shown to be involved in endothelial cell development in zebrafish (Hu et al. 2006). This expansion of HSP70-12 genes has now been identified in two oyster species and therefore would appear to represent a *Crassostrea*-specific expansion, rather than just *C. gigas*. HSP70 genes have previously shown strong correlation with chemical exposure, including hydrocarbons as used in this study, and have been suggested as candidate biomarkers for environmental toxicology (Boutet et al., 2004; Cruz-Rodríguez and Chu, 2002; Snyder et al., 2001). In these studies the inducible form of HSP70 was targeted in a candidate gene approach, as these were the only gene data available at the time, but more recent analyses have shown HSP70-12B to be induced in response to thermal challenge in *C. gigas* (Clark et al., 2013).

Table 4 should be placed here

Analysis of the transcriptome mapping shows up-regulation of HSC71, GRP78, two members of the HSP70-12A gene family and 3 members of the HSP70-12B family (Figure 2; Table 3). The variability in expression levels is expected as expansion of a gene usually results in sub-functionalization and partition of function between the different family members (Force et al., 1999). Indeed in each case of HSP70-12A and HSP70-12B, one family member shows more up regulation than the others in response to chemical challenge: contig 1465 (12A) in response to

PHE and contig 5614 (12B) in response to diesel WAF and therefore present as prime candidates for their further development as biomarkers. This expansion of members of the HSP70-12 family clearly adds a further level of complication in developing specific biomarkers for environmental monitoring in oysters, as full characterisation of all family members is needed to ensure biomarker specificity, along with the identification of which family is up-regulated in response to a particular toxin and also the time course involved, as has been identified in other stress response studies (c.f. Clark and Peck, 2009a).

In contrast to both the CYP and HSP70 families, other genes which may potentially be involved in xenobiotic metabolism do not appear to show any additional duplicated members compared with other species.

3.6. Transcripts encoding putative glutathione *S*-transferases (GSTs) and sulfotransferases

In the *C. brasiliana* dataset 26 contigs were identified with high sequence similarity to GSTs and 20 were assigned to seven cytosolic classes and six to three microsomal GSTs (Table 5). Most of the identified GSTs were assigned to the sigma (6 sequences), followed by the omega class (4 sequences), members of which are known to play a role in the xenobiotic detoxification and in the protection against oxidative stress (Board et al. 2000; Fonseca et al., 2010; Milan et al., 2011). Indeed, omega GST has been reported to be useful as a biomarker for hydrocarbon and domestic sewage exposure in oysters (Boutet et al., 2004; Lüchmann et al., 2014; Medeiros et al., 2008). The remaining contigs identified here with matches to cytosolic forms were further designated into putative pi, mu, theta and alpha classes. Of these, pi GST is known to inactivate products of oxidative damage, such as lipoperoxidation products, lipid hydroperoxides and their derivatives (Doyen et al., 2008) which has been implicated to play an important role in the detoxification of BaP in scallops (Miao et al., 2011). The mu class has also been proposed as biomarkers for hydrocarbons exposure in oysters, since it was over-expressed in *C. gigas* under chemical stress conditions (Boutet et al., 2004). Similarly, the theta class was responsive to organic compounds in the flounder *Platichthys flesus*, and has been identified as the most responsive cytosolic GST in fish challenged with organic compounds (Williams et al., 2008). This family was well represented in this transcriptome data although very little is known of this class in molluscs (Whalen et al., 2008). With regards to microsomal GSTs, six contigs were likely to be assigned to MAPEG family, with subgroups 1 (MGST1) and 3 (MGST3) most represented (Table 5). MAPEG members constitute a unique branch where most of the proteins are involved in the

production of eicosanoids (Hayes et al., 2005), although evidence shows that microsomal GSTs are capable of detoxifying organic xenobiotics in fish and human (Hayes et al., 2005; Williams et al., 2008). In general, the GSTs did not show much up-regulation in response to chemical challenge (Figure 2; Table 3). The main candidates for biomarkers are GST sigma in response to diesel WAF, GST microsomal 2 in response to sewage, whilst GST microsomal 1 showed elevated expression in response to both PHE and sewage (Figure 2; Table 3).

Table 5 should be placed here

Finally, transcripts encoding sulfotransferases (SULT) were also identified in the *C. brasiliana* transcriptome, although only one (putative *SULT1C4*) was slightly up-regulated in response to PHE only (Table 5). In rats, however, the protein identified as SULT1C1 was down-regulated following a short-term PHE exposure (Park et al., 2013). Furthermore, while a recent study has shown a clear association between increased expression of SULTs enzymes and PAHs in the channel catfish *Ictalurus punctatus* (Gaworecki et al., 2004), contradictory results in *SULT* mRNA levels have also been detected in mice treated with microsomal enzyme inducers (Alnouti and Klaasen, 2008). Thus, the SULTs inducibility is yet to be confirmed for several species, including fish and molluscs (Janer et al., 2005; Milan et al., 2011; Roméo and Wirgin, 2011). However, the presence of putative SULTs, a key component of phase II metabolism of endogenous and exogenous compounds, in *C. brasiliana* suggests a role in the detoxification or endocrine metabolism of oysters, as previously suggested for mussels (Janer et al., 2005; Lavado et al., 2006).

3.7. Transcripts encoding putative multidrug resistance proteins (MDRs)

Conjugates formed by phase II reactions are eliminated from the cells by the transport across the plasma membrane into the extracellular space, which is mediated by the multidrug resistance proteins (MDRs) (Homolya et al., 2003). Nine MDRs exist and all belong to the superfamily of ATP-Binding Cassette (ABC) transporters (Hayes et al., 2005). In mammals, MDR1 is thought to export phase II-by products and compounds complexed with endogenous glutathione (GSH), and plays, therefore, an essential role in detoxification and defence against oxidative stress (Homolya et al., 2003). In aquatic invertebrates, MDRs are called as multixenobiotic resistance (MXR), and even without a clear classification as of mammals, are known to provide protection against

toxicants (Luckenbach and Epel, 2008). In fact, previous examination has shown an over-expression of MXR-like genes in *C. gigas* exposed to domestic sewage (Medeiros et al., 2008). In this data set three members of the MDR family were identified: MDR1, 3 and 4 (Table 5), of which MDR3 presents as a good potential candidate for PHE monitoring and suggest the importance of MDRs genes in providing protection against environmental pollutants of *C. brasiliana*.

3.8. Detection of antioxidant genes of interest for ecotoxicology

Given that the exposure to pollutants and further biotransformation may generate reactive oxygen species (ROS), which create harmful environment for cellular macromolecules, we were also interested in genes that participate in general antioxidant defence mechanisms. Cellular protection against the deleterious effects of ROS cells has been attributed to a complex network composed by both enzymatic and nonenzymatic antioxidants (Halliwell and Gutteridge, 2007). A number of contigs encoding genes putatively involved in the antioxidative system were identified in the *C. brasiliana* transcriptome with the thioredoxin-related contigs most represented (Table 6). Thioredoxin-related enzymes are involved in thiol-based redox regulation and belong to the thioredoxin (Trx) superfamily, which also consists of glutathione peroxidase, peroxiredoxin and glutaredoxin gene family members (Shchedrina et al., 2007). The activities of these enzymes have been proposed to protect against metal exposure and virus-induced oxidative stress in molluscs (Nikapitiya et al., 2009; Trevisan et al., 2011; 2014), although none were significantly up-regulated in response to the treatments applied here. Contigs were also identified with high sequence similarity to superoxide dismutase (SOD) and catalase (CAT) enzymes. With regards to the SOD genes, manganese SOD (MnSOD) and copper-zinc SOD (CuZnSOD) were identified. These corresponded to the mitochondrial and cytoplasmic isoforms of enzyme as identified in mammals (Zlatkovic and Filipovic, 2011), with only the cytoplasmic form showing any up-regulation in response to chemical exposure (Table 3). The two contigs with high sequence similarity to catalase genes represented a duplication of this gene. This had been identified in another oyster species, *C. hongkongensis* in a duplication event which followed the divergence of the bivalves and gastropods (Zhang et al., 2011). In this case, both genes were up-regulated (Table 3) with the A form showing a general up-regulation in response to all three chemicals, but with the B form, more specific to PHE. Catalytic activities of SOD and CAT have been identified as being involved in the cellular protection against ROS of molluscs

(mussels and limpets) challenged with environmental pollutants and environmental stress (Ansaldo et al., 2005; Cheung et al., 2004; Lima et al., 2007; Zhang et al., 2011). However, some discrepancy between gene transcription and enzymatic activity of both antioxidants has recently been reported in *Mytilus edulis* challenged with chemical stressors (Giuliani et al., 2013). Furthermore, two contigs were identified with high sequence similarity to quinone oxidoreductase, or DT-diaphorase. The latter enzyme is generally considered as a detoxification enzyme because of its ability to reduce reactive quinones to less reactive and less toxic hydroquinones (Siegel et al., 2004). Knowledge about DT-diaphorase in invertebrates is still limited, but a few reports have demonstrated its potential function as antioxidant in molluscs (*i.e.* Manduzio et al., 2005). Altogether these findings show promising insight for understanding antioxidant metabolism in oysters, more detailed and targeted functional studies will be required to elucidate the role and regulation of such genes in *C. brasiliana.*

Table 6 should be placed here

3.9. Candidate genes for environmental ecotoxicology

This *C. brasiliana* transcriptome was generated from RNA extracted from animals exposed to three common types of pollutant (diesel WAF, PHE and sewage), which enabled the identification of transcripts potentially involved in phase I, phase II and phase III xenobiotic metabolism, besides stress and antioxidant response. Preliminary expression analysis showed that some of these transcripts were indeed up-regulated in response to chemical exposure. The results also suggest that some classes of genes are more influenced by a particular contaminant, indicating their potential application as molecular biomarkers in biomonitoring programs, but it is likely that the more accurate results will be achieved using a panel of different genes as biomarkers. This variable expression is seen in other species. For example, mussels *M. edulis* exposed to environmental pollutants showed differential regulation in the transcription level of cytochrome P450 genes, which was dependent on the chemical class, with 10% sewage extract exerting no effects on these genes whilst estrogenic compounds were negatively correlated to CYP3 levels (Cubero-Leon et al., 2012). In particular, differential expression is to be expected for those families where gene duplication events have produced numerous paralogues and retention of the extra copies is reliant upon sub-functionalisation (Force et al., 1999), as was seen with the duplicated CYP3 genes in *M. edulis* in the exposures above (Cubero-Leon et al.,

2012). It should also be noted that these were single time point and single concentration exposure experiments and more detailed experiments using approaches such as real time PCR (qPCR) and/or transcriptional profiling will be required to validate the results from this preliminary analysis. The latter method may prove particularly useful where there is an expansion in gene family members and it is not necessarily the "classical", best known family members that are responsive to the chemical or environmental challenge (c.f. Clark and Peck, 2009a; Clark et al., 2013). Also these represent laboratory-induced experiments and there is an additional requirement to link these through to responses in the natural environment, as the two types of exposure and the associated stress response may not be equivalent (c.f. Clark and Peck, 2009b). This is especially true in ecotoxicology where bioavailability of contaminants may hindering the response, mixtures of compounds may produce synergistic or antagonistic responses along with differential rates of xenobiotic metabolism depending on the toxin involved (Cruz-Rodríguez and Chu, 2002), with the latter requiring detailed time course studies. Furthermore, biotic and environmental factors such as physic-chemicals parameters are always important considerations when studying genes transcription as biomarkers. In these cases, qPCR of candidate genes may provide and relatively quick and easy, cost-effective preliminary test on which to base future studies using more expensive transcriptional profiling technologies.

4. Conclusions

This study is the first to describe a *de novo* transcriptome for the mangrove oyster *Crassostrea brasiliana*, emphasizing pathways related to detoxification of environmental contaminants. To date, the lack of genomics data available for this species has hampered characterization of the molecular mechanisms underlying resistance to aquatic contaminants and these data (circa 20,000 contigs) represent a dramatic expansion of existing cDNA sequences available for *C. brasiliana*. We have identified genes that are potential candidates as biomarkers for ecotoxicological studies including those encoding enzymes putatively involved in the biotransformation of xenobiotics and those encoding enzymes of the antioxidant defence system. Since this study was initiated, two oyster genomes have been published (*C. gigas* and *P. fucata*) (Zhang et al., 2012a; Takeuchi et al., 2012), which have aided in the annotation process. In particular our data show the same expansion of cytochrome P450 and HSP7012 gene family members as was shown in the *C. gigas* genome (Zhang et al., 2012a), indicating that the duplication events generating these particular genes may be basal to the *Crassostrea* lineage.

Why these particular gene families should be expanded in (at least two) *Crassostrea* species is unknown, but these are inter-tidal animals inhabiting environments where conditions can change rapidly and acute short term tolerance to environmental challenge is essential for survival. They would therefore be expected to be physiologically robust with an efficient "stress" response and these expanded gene sets are clearly part of their enhanced defence mechanisms. Both HSP70-12A and HSP70-12B are divergent members of the HSP70 family and it has been suggested that they are the more ancient forms (Han et al., 2003), which may partly explain why these particular genes, rather than the classical inducible forms, have massively expanded in *Crassostrea.* The data here, with the up-regulation of transcripts involved in the biotransformation of xenobiotics is in agreement with previous reports on the enzymatic activities on bivalves exposed to environmental contaminants, and adds to the current knowledge on the molecular biology and biochemistry of the stress response in oysters. Furthermore, the data generated in this study can be used as reference transcriptome for further transcriptional profiling studies on *C. brasiliana* and other bivalve species to address the molecular mechanisms underlying the susceptibility to pollutants and differential stress resistance. In the longer term the aim of this research is to develop effective biomonitoring programs using *C. brasiliana* as sentinel species in Brazilian coastal regions.

Acknowledgements

This research was supported by grants from NERC, UK to JAG and CNPq to ACDB (CT-Petro 550706/2005-4 and CNPq INCT-TA). KHL was a Guest Student at the Marine Biological Association of the United Kingdom and Glasgow Caledonian University and was supported by a CAPES Ph.D. Fellowship and CNPq Ph.D. Sandwich Fellowship, Brazil. ACDB was recipient of the CNPq Productivity Fellowship, Brazil. MSC and MAST were funded by NERC core funding to the British Antarctic Survey. We would like to thank Dr. Fabrício Flores-Nunes, Dr. Tarquin S. Dorrington and M.Sc. Christielly Rodrigues and for the assistance during experiments and Mr Jamie Oliver (British Antarctic Survey) for his help with Figure 2. We are grateful to Dr. Cláudio M.R. Melo and to M.Sc. Carlos H.A.M. Gomes for supplying the oysters used in this study.

References

Alnouti, Y., Klaassen, C.D., 2008. Regulation of sulfotransferase enzymes by prototypical

microsomal enzyme inducers in mice. J. Pharmacol. Exp. Ther. 324, 612-621.

Ansaldo, M., Najle, R., Luquet, C.M., 2005. Oxidative stress generated by diesel seawater contamination in the digestive gland of the Antarctic limpet *Nacella concinna*. Mar. Environ. Res. 59, 381-390.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. J. R. Stat. Soc. Series B, 57, 289-300.

Bettencourt, R., Pinheiro, M., Egas, C., Gomes, P., Afonso, M., Shank, T., Santos, R.S., 2010. High-throughput sequencing and analysis of the gill tissue transcriptome from the deep-sea hydrothermal vent mussel *Bathymodiolus azoricus*. BMC Genom. 11, 559.

Board, P.G., Coggan, M., Chelvanayagam, G., Easteal, S., Jermiin, L.S., Schulte, G.K., Danley, D.E., Hoth, L.R., Griffor, M.C., Kamath, A.V., Rosner, M.H., Chrunyk, B.A., Perregaux, D.E., Gabel, C.A., Geoghegan, K.F., Pandit, J., 2000. Identification, characterization, and crystal structure of the omega class glutathione transferases. J. Biol. Chem. 275, 24798-24806.

Boutet, I., Tanguy, A., Moraga, D., 2004. Characterisation and expression of four mRNA sequences encoding glutathione S-transferases pi, mu, omega and sigma classes in the Pacific oyster *Crassostrea gigas* exposed to hydrocarbons and pesticides. Mar. Biol. 146, 53-64.

Cheung, C.C.C., Siu, W.H.L., Richardson, B.J., De Luca-Abbott, S.B., Lam, P.K.S., 2004. Antioxidant responses to benzo[a]pyrene and Aroclor 1254 exposure in the green-lipped mussel, *Perna viridis*. Environ. Pollut. 128, 393-403.

Clark, M.S., Peck, L.S., 2009a. HSP70 heat shock proteins and environmental stress in Antarctic marine organisms: A mini-review. Mar. Gen. 2, 11-18.

Clark, M.S., Peck, L.S., 2009b. Triggers of the HSP70 stress response: environmental responses and laboratory manipulation in an Antarctic marine invertebrate (*Nacella concinna*). Cell Stress Chaperon. 14, 649-660.

Clark, M.S., Thorne, M.A., Vieira, F.A., Cardoso, J.C., Power, D.M., Peck, L.S., 2010. Insights into shell deposition in the Antarctic bivalve *Laternula elliptica*: gene discovery in the mantle transcriptome using 454 pyrosequencing. BMC Genom. 11, 362.

Clark, M.S., Thorne, M.A.S., Toullec, J.-Y., Meng, Y., Guan, L.L., Peck, L.S., Moore, S., 2011. Antarctic Krill 454 Pyrosequencing Reveals Chaperone and Stress Transcriptome. PLoS One 6, e15919.

Clark, M.S., Thorne, M.A.S., Amaral, A., Viera, F., Batista, F.M., Reis, J., Power, D.M, 2013. Identification of molecular and physiological responses to chronic environmental challenge in an invasive species: the Pacific oyster, *Crassostrea gigas.* Ecol. Evol. 3, 3283-3297.

Conesa, A., Götz, S., García-Gómez, J.M., Terol, J., Talón, M., Robles, M., 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21, 3674-3676.

Craft, J.A., Gilbert, J.A., Temperton, B., Dempsey, K.E., Ashelford, K., Tiwari, B., Hutchinson, T.H., Chipman, J.K., 2010. Pyrosequencing of *Mytilus galloprovincialis* cDNAs: tissue-specific expression patterns. PLoS One 5, e8875.

Cruz-Rodríguez, L.A., Chu, F.-L.E., 2002. Heat shock protein (HSP70) response in the eastern oyster, *Crassostrea virginica,* exposed to PAHs sorbed to suspended artificial clay particles and to suspended field contaminated sediments. Aquat. Toxicol. 60, 157-168.

Cubero-Leon, E., Puinean, A.M., Labadie, P., Ciocan, C., Itoh, N., Kishida, M., Osada, M., Minier, C., Hill, E.M., Rotchell, J.M., 2012. Two CYP3A-like genes in the marine mussel *Mytilus edulis*: mRNA expression modulation following short-term expo- sure to endocrine disruptors. Mar. Environ. Res. 74, 32-39.

Cui, Z.M., Wang, P., Sun, L., Liu, H., Yang, J., Li, X., Kang, L., Huang, Y. et al., 2010. Lipopolysaccharide-evoked HSPA12B expression by activation of MAPK cascade in microglial cells of the spinal cord. J. Neurol. Sci. 294, 29-37.

Doyen, P., Bigot, A., Vasseur, P., Rodius, F., 2008. Molecular cloning and expression study of pi-class glutathione S-transferase (pi-GST) and selenium-dependent glutathione peroxidase (Se-GPx) transcripts in the freshwater bivalve *Dreissena polymorpha*. Comp. Biochem. Physiol. C 147, 69-77.

Du, H., Bao, Z., Hou, R., Wang, S., Su, H., Yan, J., Tian, M., Li, Y. et al., 2012. Transcriptome sequencing and characterisation for the sea cucumber *Apostichopus japonicas* (Selenka 1867). PLoS One 7, e33311.

Feder, M.E., Hofmann, G.E., 1999. Heat shock proteins, molecular chaperones and their stress response: Evolutionary and Ecological Physiology. Annu. Rev. Physiol. 61, 243-282.

Feyereisen, R., 2006. Evolution of insect P450. Biochem. Soc. Trans. 34, 1252-1255.

Fink, A.L., 1999. Chaperone-mediated protein folding. Physiol. Rev. 79, 425-449.

Fleury, E., Huvet, A., Lelong, C., Lorgeril, J., Boulo, V., Gueguen, Y., Bachère, E. et al., 2009. Generation and analysis of a 29,745 unique Expressed Sequence Tags from the Pacific oyster

(*Crassostrea gigas*) assembled into a publicly accessible database: the GigasDatabase. BMC Genom. 10, 341.

Fonseca, R.R., Johnson, W.E., O'Brien, S.J., Vasconcelos, V., Antunes, A., 2010. Molecular evolution and the role of oxidative stress in the expansion and functional diversification of cytosolic glutathione transferases. BMC Evol. Biol. 10, 281.

Forbes, V.E., Palmqvist, A., Bach, L., 2006. The use and misuse of biomarkers in ecotoxicology. Environ. Toxicol. Chem. 25, 272-280.

Force, A., Lynch, M., Pickett, F.B., Amores, A., Yan, Y.L., Postlethwait, J., 1999. Preservation of duplicate genes by complementary, degenerative mutations. Genetics 151, 1531-1545.

Fraser, B.A., Weadick, C.J., Janowitz, I., Rodd, H., Hughes, K.A., 2011. Sequencing and characterization of the guppy (*Poecilia reticulata*) transcriptome. BMC Genom. 12, 202.

Gaworecki, K.M., Rice, C.D., Van den Hurk, P., 2004. Induction of phenol-type sulfotransferase and glucuronosyltransferase in channel catfish and mummichog. Mar. Environ. Res. 58, 525-528.

Giuliani, M.E., Benedetti, M., Arukwe, A., Regoli, F., 2013. Transcriptional and catalytic responses of antioxidant and biotransformation pathways in mussels, *Mytilus galloprovincialis*, exposed to chemical mixtures. Aquat. Toxicol. 134-135, 120-127.

Goldstone, H.M., Stegeman, J.J., 2006. A revised evolutionary history of the CYP1A subfamily: gene duplication, gene conversion, and positive selection. J. Mol. Evol. 62, 708-717.

Gross, M., 2004. Emergency services: A bird's eye perspective on the many different functions of stress proteins. Curr. Protein Peptide Sci. 5, 213-223.

Guo, H., Bao, Z., Du, H., Zhang, L., Wang, S., Sun, L., Mou, X., Hu, X., 2013. Identification of Cytochrome P450 (CYP) Genes in Zhikong Scallop (*Chlamys farreri*). J. Ocean Univ. China 12, 97-102.

Halliwell, B., Gutteridge, J., 2007. Free Radicals in Biology and Medicine. Oxford University Press, Nova York.

Han, Z., Truong, Q.A., Park, S., Breslow, J.L., 2003. Two HSP70 family members expressed in atherosclerotic lesions. Proc. Natl. Acad. Sci. USA 100, 1256-1261.

Hannam, M.L., Bamber, S.D., Galloway, T.S., Moody, A.J., Jones, M.B., 2010. Effects of the model PAH phenanthrene on immune function and oxidative stress in the haemolymph of the temperate scallop *Pecten maximus.* Chemosphere 78, 779-784.

Hao, D.C., Ge, G., Xiao, P., Zhang, Y., Yang, L., 2011. The first insight into the tissue specific *Taxus* transcriptome via Illumina second generation sequencing. PLoS One 6, e21220.

Hartl, F.U., 1996. Molecular chaperones in cellular protein folding. Nature 381, 571-580.

Hartl, F.U., Hartl-Meyer, M., 2002. Molecular chaperones in the cytosol: from nascent chain to folded protein. Science 295, 1852-1858.

Hayes, J.D., Flanagan, J.U., Jowsey, I.R., 2005. Glutathione transferases. Annu. Rev. Pharmacol. Toxicol. 45, 51-88.

Homolya, L., Varadi, A., Sarkadi, B., 2003. Multidrug resistance-associated proteins: Export pumps for conjugates with glutathione, glucuronate or sulfate. Biofactors 17, 103-114.

Hou, R., Bao, Z., Wang, S., Su, H., Li, Y., Du, H., Hu, J., Wang, S., Hu, X., 2011. Transcriptome Sequencing and *De Novo* Analysis for Yesso Scallop (*Patinopecten yessoensis*) Using 454 GS FLX. PLoS One 6, e21560.

Hu, G., Tang, J., Zhang, B., Lin, Y., Hanai, J.-I., Galloway, J., Bedell, V., Bahary, N. et al., 2006. A novel endothelial-specific heat shock protein HspA12B is required in both zebrafish development and endothelial functions *in vitro*. J. Cell Sci. 119, 4117-4126.

Janer, G., Mesia-Vela, S., Kauffman, F.C., Porte, C., 2005. Sulfatase activity in the oyster *Crassostrea virginica*: Its potential interference with sulfotransferase determination. Aquat. Toxicol. 15, 92-95.

Joubert, C., Piquemal, D., Marie, B., Manchon, L., Pierrat, F., Zanella-Cleon, I., Cochennec-Laureau, N., Gueguen, Y., Montagnani, C., 2010. Transcriptome and proteome analysis of *Pinctada margaritifera* calcifying mantle and shell: focus on biomineralization. BMC Genom. 11, 613.

Karatolos, N., Pauchet, Y., Wilkinson, P., Chauhan, R., Denholm, I., Gorman, K., Nelson, D.R., Bass, C., Ffrench-Constant, R.H., Williamson, M.S., 2011. Pyrosequencing the transcriptome of the greenhouse whitefly, *Trialeurodes vaporariorum* reveals multiple transcripts encoding insecticide targets and detoxifying enzymes. BMC Genom. 12, 56.

Kikuta, Y., Kusunose, E., Kusunose, M., 2002. Prostaglandin and leukotriene omega-hydroxylases. Prostaglandins Other Lipid Mediat. 68-69, 345-362.

Kubota, A., Stegeman, J.J., Goldstone, J.V., Nelson, D.R., Kim, E.Y., Tanabe, S., Iwata, H., 2011. Cytochrome P450 CYP2 genes in the common cormorant: Evolutionary relationships with 130 diapsid CYP2 clan sequences and chemical effects on their expression. Comp. Biochem. Physiol. C Toxicol. Pharmacol. 153, 280-289.

Lavado, R., Janer, G., Porte, C., 2006. Steroid levels and steroid metabolism in the mussel *Mytilus edulis*: the modulating effect of dispersed crude oil and alkylphenols. Aquat. Toxicol. 78, 65-72.

Lazoski, C., Gusmão, J., Boudry, P., Solé-Cava, A.M., 2011. Phylogeny and phylogeography of Atlantic oyster species: evolutionary history, limited genetic connectivity and isolation by distance. Mar. Ecol. Prog. Ser. 426, 197-212.

Lima, I., Moreira, S.M., Rendón-Von Osten, J., Soares, A.M.V.M., Guilhermino, L., 2007. Biochemical responses of the marine mussel *Mytilus galloprovincialis* to petrochemical environmental contamination along the North-western coast of Portugal. Chemosphere 66, 1230-1242.

Lüchmann, K.H., Mattos, J.J., Siebert, M.N., Granucci, N., Dorrington, T.S., Bicego, M.C., Taniguchi, S., Sasaki, S.T., Daura-Jorge, F.G., Bainy, A.C.D., 2011. Biochemical biomarkers and hydrocarbons concentrations in the mangrove oyster *Crassostrea brasiliana* following exposure to diesel fuel water-accommodated fraction. Aquat. Toxicol. 105, 652-660.

Lüchmann, K.H., Dafre, A.L., Trevisan, R., Craft, J.A., Meng, X., Mattos, J.J., Zacchi, F.L., Dorrington, T.S., Schroeder, D.C., Bainy, A.C.D., 2014. A light in the darkness: New biotransformation genes, antioxidant parameters and tissue-specific responses in oysters exposed to phenanthrene. Aquat. Toxicol. 162, 324-334.

Luckenbach, T., Epel, D., 2008. ABCB- and ABCC-type transporters confer multixenobiotic resistance and form an environment-tissue barrier in bivalve gills. Am. J. Physiol. Regul. Integr. Comp. Physiol. 294, R1919-R1929.

Manduzio, H., Rocher, B., Durand, F., Galap, C., Leboulenger, F., 2005. The point about oxidative stress in molluscs. ISJ 2, 91-104.

Medeiros, I.D., Siebert, M.N., Toledo-Silva, G., Rodrigues, T.B., Marques, M.R.F., Bainy, A.C.D., 2008. Induced gene expression in oyster *Crassostrea gigas* exposed to sewage. Environ. Toxicol. Pharmacol. 26, 362-365.

Meyer, E., Aglyamova, G.V., Wang, S., Buchanan-Carter, J., Abrego, D., Colbourne, J.K., Willis, B.L., Matz, M.V., 2009. Sequencing and de novo analysis of a coral larval transcriptome using 454 GSFlx. BMC Genom. 10, 219.

Miao, J., Pan, L., Liu, N., Xu, C., Zhang, L., 2011. Molecular cloning of CYP4 and GSTpi homologues in the scallop *Chlamys farreri* and its expression in response to Benzo[a]pyrene exposure. Mar. Genom. 4, 99-108.

Milan, M., Coppe, A., Reinhardt, R., Cancela, L.M., Leite, R.B., Saavedra, C., Ciofi, C., Chelazzi, G., Patarnello, T., Bortoluzzi, S., Bargelloni, L., 2011. Transcriptome sequencing and microarray

development for the Manila clam, *Ruditapes philippinarum*: genomic tools for environmental monitoring. BMC Genom. 12, 234.

Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A.C., Kanehisa, M., 2007. KAAS: an automatic genome annotation and pathway reconstruction server. Nucleic Acids Res. 35, 182-185.

Nebert, D.W., Russell, D.W., 2002. Clinical importance of the cytochromes P450. Lancet 360, 1155-1162.

Nelson, D.R., Goldstone, J.V., Stegeman, J.J., 2013. The cytochrome P450 genesis locus: the origin and evolution of animal cytochrome P450s. Phil. Trans. R. Soc. B, 368, 20120474.

Nikapitiya, C., De Zoysa, M., Whang, I., Kim, C.G., Lee, Y.H., Kim, S.J., Lee, J., 2009. Molecular cloning, characterization and expression analysis of peroxiredoxin 6 from disk abalone *Haliotis discus discus* and the antioxidant activity of its recombinant protein. Fish. Shellfish Immunol. 27, 239-249.

Niu, D., Wang, L., Sun, F., Liu, Z., Li, J., 2013. Development of molecular resources for an intertidal clam, *Sinonovacula constricta*, using 454 transcriptome sequencing. PLoS One 8, e67456.

Oliveira, M., Pacheco, M., Santos, M.A., 2007. Cytochrome P4501A, genotoxic and stress responses in golden grey mullet (*Liza aurata*) following short-term exposure to phenanthrene. Chemosphere 66, 1284-1291.

Parchman, T.L., Geist, K.S., Grahnen, J.A., Benkman, C.W., Buerkle, C.A., 2010. Transcriptome sequencing in an ecologically important tree species: assembly, annotation, and marker discovery. BMC Genom. 11, 180.

Park, S.-K., Ryu, A.-R., Lee, M.-Y., 2013. Protein expression profiling in the liver of rats exposed to phenanthrene. Mol. Cell Toxicol. 9, 267-275.

Parsell, D.A., Lindquist, S., 1993. The function of heat-shock proteins in stress tolerance – degradation and reactivation of damaged proteins. Ann. Rev. Genet. 27, 437-496.

Putnam, N.H., Srivastavia, M., Hellsten, U., Dirks, B., Chapman, J., Salamov, A., Terry, A., Shaprio, H. et al., 2007. Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organisation. Science 317, 86-94.

Rewitz, K.F., Styrishave, B., Løbner-Olesen, A., Anderson, O., 2006. Marine invertebrate cytochrome P450: emerging insights from vertebrate and insect analogies. Comp. Biochem. Physiol. C 143, 363-381.

Roméo, M., Wirgin, I., 2011. Biotransformation of organic contaminants and the acquisition of resistence, in: Amiard-Triquet, C., Rainbow, P.S., Romeo, M. (Eds.), Tolerance to Environmental Contaminants. CRC Press, Florida. pp. 175-208.

Schober, W., Pusch, G., Oeder, S., Reindl, H., Behrendt, H., et al., 2010. Metabolic activation of phenanthrene by human and mouse cytochromes P450 and pharmacokinetics in CYP1A2 knockout mice. Chemico-Biol. Inter. 183, 57-66.

Shchedrina, V.A., Novoselov, S.V., Malinouski, M.Y., Gladyshev, V.N., 2007. Identification and characterization of a selenoprotein family containing a diselenide bond in a redox motif. Proc. Natl. Acad. Sci. USA 104, 13919-13924.

Siegel, D., Gustafson, D.L., Dehn, D.L., Han, J.Y., Boonchoong, P., Berliner, L.J., Ross, D., 2004. NAD(P)H:quinone oxidoreductase 1: role as a superoxide scavenger. Mol. Pharmacol. 65, 1238-1247.

Simpson, A.E.C.M., 1997. The cytochrome P450 4 (CYP4) family. Gen. Pharmac. 28, 351-359.

Singer, M.M., Aurand, D., Bragin, G.E., Clark, J.R., Coelho, G.M., Sowby, M.L., Tjeerdema, R.S., 2000. Standardization of the preparation and quantitation of water-accommodated fractions of petroleum for toxicity testing. Mar. Pollut. Bull. 40, 1007-1016.

Snyder, M., Girvetz, E., Mulder, E., 2001. Induction of marine mollusc stress proteins by chemical or physical stress. Arch. Environ. Con. Tox. 44, 22-29.

Sodergren, E., Weinstock, G.M., Davidson, E.H., Cameron, R.A., Gibbs, R.A., Angerer, R.C., Angerer, L.M., Anone, M.I. et al., 2006. The genome of the sea urchin *Stronglyocentrotus purpuratus*. Science 314, 941-952.

Takeuchi, T., Kawashima, T., Koyanagi, R., Gyoja, F., Tanaka, M., Ikuta, T., Shoguchi, E., Fujiwara, M. et al., 2012. Draft genome of the pearl oyster *Pinctada fucata*: a platform for understanding bivalve biology. DNA Res. 19, 117-130.

Tanguy, A., Bierne, N., Saavedra, C., Pina, B., Bachere, E., Kube, M., Bazin, E., Bonhomme, F., Boudry, P., Boulo, V., et al., 2008. Increasing genomic information in bivalves through new EST collections in four species: development of new genetic markers for environmental studies and genome evolution. Gene 408, 27-36.

Toledo-Silva, G., Siebert, M.N., Medeiros, I., Sincero, T., Moraes, M., Goldstone, J.V., Bainy, A.C.D., 2008. Cloning a new cytochrome P450 isoform (CYP356A1) from oyster *Crassostrea gigas*. Mar. Environ. Res. 66, 15-18.

Trevisan, R., Mello, D.F., Fisher, A.S., Schuwerack, P.M., Dafre, A.L., Moody, A.J., 2011. Selenium in water enhances antioxidant defenses and protects against copper-induced DNA damage in the blue mussel *Mytilus edulis*. Aquat. Toxicol. 101, 64-71.

Trevisan, R., Delapedra, G., Mello, D.F., Arl, M., Schmidt, E.C., Meder, F., Monopoli, M., Cargnin-Ferreira, E., Bouzon, Z.L., Fisher, A.S., Sheehan, D., Dafre, A.L., 2014. Gills are an initial target of zinc oxide nanoparticles in oysters *Crassostrea gigas*, leading to mitochondrial disruption and oxidative stress. Aquat. Toxicol. 153, 27-38.

US EPA: *United States Environmental Protection Agency Priority Pollutants*; 2009. http://www.epa.gov/waterscience/methods/pollutants.htm (Accessed Feb/2015).

Walker, C.H., Sibly, R.M., Hopkin, S.P., Peakall, D.B., 1996. Principles of Ecotoxicology. Taylor & Francis, London.

Wang, Y., Guo, X., 2007. Development and Characterization of EST-SSR Markers in the Eastern Oyster *Crassostrea virginica.* Mar. Biotechnol. 9, 500-511.

Whalen, K.E., Morin, D., Lin, C.Y., Tjeerdema, R.S., Goldstone, J.V., Hahn, M.E., 2008. Proteomic identification, cDNA cloning and enzymatic activity of glutathione *S*-transferases from the generalist marine gastropod, *Cyphoma gibbosum*. Arch. Biochem. Biophys. 478, 7-17.

Williams, T.D., Diab, A., Ortega, F., Sabine, V.S., Godfrey, R.E., et al., 2008. Transcriptomic responses of European flounder (*Platichthys flesus*) to model toxicants. Aquat. Toxicol. 90, 83-91.

Zanette, J., Goldstone, J.V., Bainy, A.C.D., Stegeman, J.J., 2010. New cytochrome P450 genes in bivalves -Identification of CYPs from mussel and oyster ESTs. Mar. Environ. Res. 69, S1-S3.

Zanette, J., Jenny, M.J., Goldstone, J.V., Parente, T., Woodin, B.R., Bainy, A.C.D., Stegeman, J.J., 2013. Identification and expression of multiple CYP1-like and CYP3-like genes in the bivalve mollusk *Mytilus edulis*. Aquat. Toxicol. 128-129, 101-112.

Zhang, Y., Fu, D., Yu, F., Liu, Q., Yu, Z., 2011. Two catalase homologs are involved in host protection against bacterial infection and oxidative stress in *Crassostrea hongkongensis*. Fish Shellfish Res. 31, 894-903.

Zhang, G., Fang, X., Guo, X., Li, L., Luo, R., Xu, F., Yang, P., Zhang, L. et al., 2012a. The oyster genome reveals stress adaptation and complexity of shell formation. Nature 490, 49-54.

Zhang, L., Gan, J., Ke, C., Liu, X., Zhao, J., You, L., Yu, J., Wu, H., 2012b. Identification and expression profile of a new cytochrome P450 isoform (CYP414A1) in the hepatopancreas of *Venerupis* (Ruditapes) *philippinarum* exposed to benzo[a]pyrene, cadmium and copper. Environ. Toxicol. Pharmacol. 33, 85-91.

Zlatkovic, J., Filipovic, D., 2011. Stress-induced alternations in CuZnSOD and MnSOD activity in cellular compartments of rat liver. Mol. Cell Biochem. 357, 143-150.

Candidate genes	Number	Number	Average contig	Average reads

Table Captions

Table 1. 454 summary statistics for for the *C. brasiliana* transcriptome assembly.

Assembly data	
Total number of reads	399,291
Total number of assembled reads	246,514
Total number of contigs	20,938
Total number of singletons	152,777
rRNA transcripts excluded	127
mtDNA transcripts excluded	90
Prokaryotic and virus transcripts excluded	65
Final number of contigs	**20,656**
Contig data	
Average contig size (bp)	575
Largest contig (bp)	4,462
Number of contigs longer than 1kb	1,556 (15.2%)
Number of contigs longer than 500bp	10,260 (49%)
N50 (bp)	911
Average number of reads per contig	12

Table 2. Summary information for the Cytochrome P450 transcripts identified in *C. brasiliana* that are putatively involved in phase I of the metabolism of biotransformation of xenobiotics. Classification of clans according to Nelson (1998).

	of contigs	of genes	size (bp)	per contigs (bp)
Clan 2				
CYP450, family 1	4	2	874	11
CYP450, family 2	14	7	767	15
CYP450, family 17	4	3	1,531	14
Clan3				
CYP450, family 3	7	6	770	5
Clan 4				
CYP450, family 4	9	6	507	2
Clan 7				
CYP450, family 7	1	1	569	2
Clan mitochondrial				
CYP450, family 12	2	2	756	3
CYP450, family 13	1	1	489	2
CYP450, family 24	1	1	916	3

Table3. Gene ratios of expression levels in *C. brasiliana* in response to the different chemical exposures. Where putative multiple forms of the genes exist, these have not been characterised, but are denoted by a number in () to differentiate them in these results.

Candidate gene	Contig	Diesel WAF	PHE	Sewage
CYP450s				
CYP1A	4236	8.75	4.21	0.87
CYP2C	762	5.84	0.62	3.77
CYP2D	14324	4.38	0.00	0.00
CYP17	1780	2.19	9.82	15.67
CYP3A	1765	4.38	7.02	3.48
Heat shock proteins				
HSC71	866	2.19	2.11	0.87
GRP78	1750	2.81	2.41	2.36
HSP7012A (1)	1465	2.19	4.21	2.61
HSP7012A (2)	5468	0.00	0.00	2.18
HSP7012B (1)	5007	2.19	1.40	2.61
HSP7012B (2)	5614	6.57	1.40	0.00
HSP7012B (3)	16631	2.19	0.00	0.87
Glutathione-S-transferases				
GST	5914	0.00	2.81	0.00
GST A	504	0.67	0.67	2.08
GST sigma (1)	7984	6.57	1.40	0.00
GST sigma (2)	9071	0.00	0.00	2.61
GST mu	1576	1.09	2.81	3.05
GST microsomal 1	1539	0.00	4.21	6.96
GST microsomal 2	3082	0.00	0.70	4.35

GST microsomal 3	1452	1.46	2.81	0.87
Sulfotransferases				
Sulfotransferase 1C4	17609	0.00	2.81	0.00
Multidrug resistance proteins				
MDR-associated protein 1	13538	0.00	0.00	2.61
MDR3	2495	0.00	7.02	0.87
Other genes of interest				
MnSOD	1433	0.36	0.70	2.90
CuZnSOD	6196	2.19	5.61	4.35
Catalase (A)	1078	4.92	4.56	2.83
Catalase (B)	12078	0.00	2.81	0.87

Table 4. Summary information for the HSP70 transcripts in *C. brasiliana.*

Candidate genes	Number of contigs	Number of genes	Average contig size (bp)	Average reads per contig
HSP70	3	1	542	4
HSC71	2	1	1652	101
HSP70 B2	5	1	838	12
HSP70 mt	2	1	503	4
GRP78	2	1	1340	25
68kDa	1	1	463	2
HSP70-12A	19	13	548	5
HSP70-12B	17	8	566	5

Table 5. Summary information for the identified *C. brasiliana* genes putatively involved in phases II and III of the metabolism of biotransformation of xenobiotics.

Candidate genes	Number of contigs	Average contig size (bp)	Average reads per contig
Phase II - Glutathione *S*-transferases			
GST alpha	2	346	2
GST omega	4	646	7
GST pi	2	819	11
GST theta	2	678	3
GST sigma	6	711	8
GST mu	1	1048	14
GST A	3	798	41
GST microsomal 1	2	541	11
GST microsomal 2	1	464	12
GST microsomal 3	3	821	14
Phase II - Sulfotransferases			
Sulfotransferases	7	586	3
Phase III - Multidrug resistance proteins			
MDR1	6	467	4
MDR3	2	655	5
MDR4	1	489	3

Table 6. Summary information for the *C. brasiliana* transcripts putatively involved in the antioxidant defence system.

Candidate genes	Number of contigs	Average contig size (bp)	Average reads per contigs
CuZn Superoxide dismutase	2	805	53
Mn Superoxide dismutase	1	1073	30
Catalase	2	1325	22
Glutathione peroxidase	2	577	7
Glutathione reductase	1	536	5
Peroxiredoxin 5	2	737	10
Peroxiredoxin 6	1	1013	13
Glutaredoxin 2	2	468	4
Glutaredoxin 5	1	390	12
Glutaredoxin C6	1	414	5

Thioredoxin	8	657	7
Thioredoxin reductase	2	564	4
Thioredoxin peroxidase	1	989	10
Quinone oxidoreductase	2	529	2
Glucose-6-phosphate dehydrogenase	2	888	5

Figure Captions

Figure 1. General enzyme classification based on enzyme commission (EC) numbers for the contigs of *C. brasiliana*.

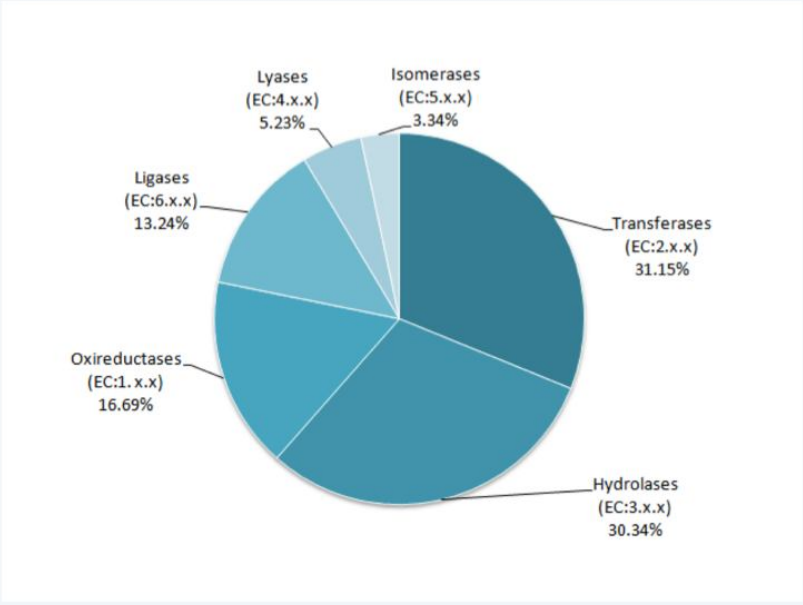
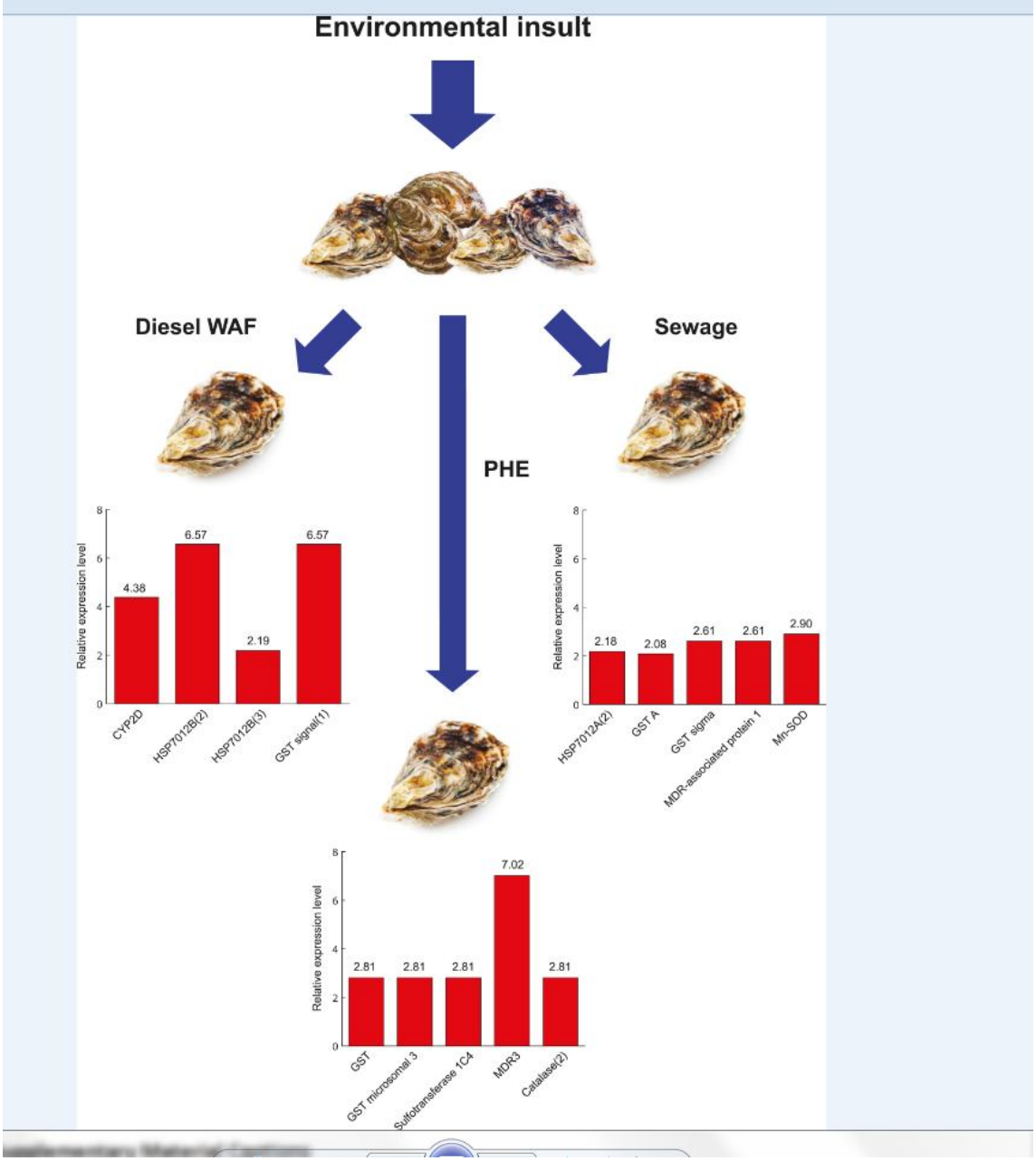


Figure 2. Schematic of expression profiling data indicating putative unique biomarkers for the environmental toxicants under study.



Supplementary Material Captions

Table S1. InterPro and Pfam Identifiers used in transcriptome searches for specific gene families.

Figure S1. Overview of the *C. brasiliana* transcriptome assembly. (A) Frequency distribution of contig lengths. (B) Frequency distribution of the number of reads assembled into contigs.